



Aalborg Universitet

**AALBORG UNIVERSITY**  
DENMARK

## **Sound localization with binaural recordings made with artificial heads**

Minnaar, Pauli; Olesen, Søren Krarup; Christensen, Flemming; Møller, Henrik

*Published in:*  
Proceedings of ICA2004,

*Publication date:*  
2004

[Link to publication from Aalborg University](#)

### *Citation for published version (APA):*

Minnaar, P., Olesen, S. K., Christensen, F., & Møller, H. (2004). Sound localization with binaural recordings made with artificial heads. In *Proceedings of ICA2004*, (pp. 3651-3654)

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### **Take down policy**

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# Sound localisation with binaural recordings made with artificial heads

*P. Minnaar, S. K. Olesen, F. Christensen, H. Møller*

Department of Acoustics, Aalborg University  
Fredrik Bajers Vej 7 B5, DK-9220 Aalborg, Denmark  
acoustics@acoustics.auc.dk

## Abstract

A series of experiments on localisation with binaural recordings made with artificial heads, done at Aalborg University, has been reviewed. The initial experiments indicated that recording with artificial heads generally is inferior to recording in the ears of selected humans. Since that time new artificial heads have been developed and existing ones may have been improved. Therefore, another experiment was done, employing recordings from 7 artificial heads and 2 selected human heads. The results from this experiment show that artificial heads still are not as good for recording as a well-selected human head – although some of the new heads come close. The accumulated results of five experiments show that there are significant differences between currently available artificial heads.

## 1. Introduction

If binaural recordings are to find widespread application a well designed artificial head is needed. This has been recognized by several companies and research institutions that have developed artificial heads. The heads must be evaluated in listening experiments, in which recordings are made and played back under strictly controlled circumstances. The heads should not only be compared to each other but also to human heads, and to real life.

A substantial amount of literature exists on localization with binaural recordings. Thorough reviews of 5 studies with recordings from human heads and of 18 studies with recordings from artificial heads are presented in Møller et al. [1] and [2] respectively.

## 2. Heads included in the study

In the current work a listening experiment was done in order to compare the performance of 7 artificial heads to each other and to 2 good human heads selected from Møller et al. [3]. The experiment has been described in detail by Minnaar et al. [4] and only the essential information are repeated here. Some of the artificial heads included in the study are commercially available, while others only exist in a few samples used in

research laboratories. These heads are briefly described here.

### 2.1. Knowles Electronics Inc. - KEMAR

The Knowles Electronics Manikin for Acoustic Research (KEMAR) is a well-known artificial head, much used in research. It has ear simulators according to IEC 711 [5] and ANSI S3.25 [6], and it conforms with geometrical and acoustical requirements of IEC 959 [7] as well as geometrical and acoustical (sound pick-up) requirements of ITU-T P.58 [8]. In the experiment reported by Møller et al. [2] the manikin was tested with 4 different pinnae (called KEMAR 1-4). Since KEMAR 2 (with DB065/066 pinnae) is the most common, it was employed in the current experiment. Brüel & Kjær 4158 and 4159 ear simulators with preamplifiers (for right and left ear respectively) and a Brüel & Kjær Nexus 2690 2-channel conditioning amplifier were used.

### 2.2. Georg Neumann GmbH - KU100

The KU100 artificial head from Neumann, like its predecessors KU80 and KU81, finds most of its application in the recording industry. It is the only head in this investigation without a torso. It has a fixed set of pinnae and comes with built-in microphones and preamplifiers. The microphone signals were fed into a 2-channel Rostec LMA 4 amplifier to obtain a line level signal.

### 2.3. Brüel & Kjær A/S - 4100

The Brüel & Kjær 4100 artificial head has the same external geometric shape as the Brüel & Kjær 4128 and 5930 tested in previous experiments [2]. However, the ear canals, microphones and preamplifiers are different, and the 4100 has a jacket and an adjustable neck ring. During recording the jacket was used and the neck was in an upright position. The built-in microphones and preamplifiers were used with a Brüel & Kjær Nexus 2690 conditioning amplifier. In the 4128 version, the Brüel & Kjær head has ear simulators according to IEC 711 and ANSI S3.25, and it conforms with the acoustical requirements of IEC 959 (but not the geometrical), and the geometrical and acoustical requirements of ITU-T P.58.



#### 2.4. Head Acoustics GmbH - HMS II

The Head Acoustics artificial head that forms part of the HMS II measurement system is well known in the engineering community. It has a stylised (mathematically describable) head and pinnae. In the version used, the HMS II has ear simulators according to IEC 711 and ANSI S3.25, and it conforms with the acoustical requirements of IEC 959 (but not the geometrical), and the geometrical and acoustical (sound pick-up) requirements of ITU-T P.58. The signals from the built-in microphones and preamplifiers were fed into a Brüel & Kjær Nexus 2690 conditioning amplifier. The system has an option of equalization during recording which was not used.

#### 2.5. Cortex Electronic GmbH - MK1

The MK1 artificial head from Cortex Electronic has an articulated neck and hips. The external shape of the head, torso and pinnae follow the geometrical descriptions in IEC 959 and ITU-T P.58, but since the head does not have ear simulators, it does not conform with the remaining requirements of the documents. The MK1 has built-in microphones, amplifiers and A/D converters, and provides a digital signal for recording. Equalization for a specific Sennheiser HE 60 electrostatic headphone, supplied by the manufacturer, was used during recording, implying that recordings were pre-equalized for this headphone.

#### 2.6. Aachen University - ITA

The artificial head developed at the Institute of Technical Acoustics at Aachen University in Germany is denoted as ITA here. It has a hard plastic head and shoulders and human-like pinnae. The built-in microphones and amplifiers were used to obtain a line-level signal. The built-in equalization and A/D converters were not employed during recording.

#### 2.7. Aalborg University - VALDEMAR

The artificial head developed at our own laboratory is named VALDEMAR after the Danish inventor Valdemar Poulsen who invented and patented the magnetic recording principle in 1898. The head and the torso have been designed from acoustical measurements on 40 humans and from anatomical data. The pinnae are casts of a human pinna (subject DOL, included in this study). Small electret microphones (Sennheiser KE 4-211-2) were inserted into earplugs with the diaphragms facing outward, and mounted flush with the ear canal entrances. A custom-made preamplifier was used.

#### 2.8. Human heads AVH and DOL

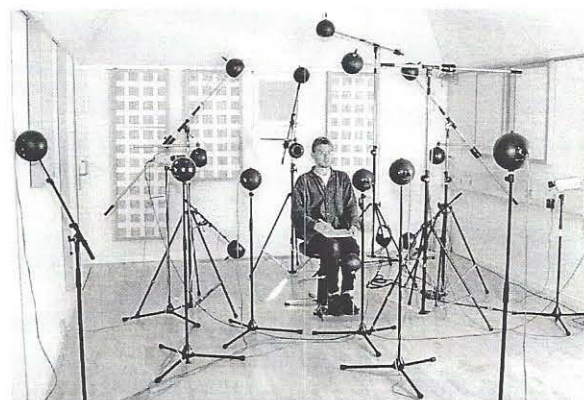
The two people used as recording heads are denoted AVH and DOL. In the study by Møller et al. [3], where 20 people listened to recordings made with 30 people's heads, recordings of AVH gave the lowest number of

median-plane errors, while recordings from DOL ranked 4th. In an inspection of acoustical measurements on 40 human pinnae, DOL's pinna was found to best represent characteristics of human pinnae. As in the case of VALDEMAR, Sennheiser KE 4-211-2 microphones were mounted flush with the blocked ear canal entrances. AVH and DOL looked straight ahead and sat perfectly still during the recording

### 3. Methods

The experiments were carried out in a listening room, where 19 loudspeakers were located around the listener. Short segments of speech or noise were presented to the listener either directly from the loudspeakers or indirectly as a binaural recording made in the same setup and reproduced by means of headphones. In both cases the loudspeakers were visible to the listener who had to keep the head still during sound presentation.

The task of the listener was to identify the loudspeaker from which he/she perceived the sound. The experiments, therefore, did not aim to measure absolute localization judgement, but rather the ability to identify a sound source in a 19-alternative forced choice task. Since the experiments were done in a room (as opposed to an anechoic environment), the reflections from the room boundaries particular to each loudspeaker contributed to the localization. The setup of the loudspeakers around a listener is shown in Figure 1.



**Figure 1.** Experimental setup of loudspeakers in a standard listening room.

In order to implement the sound presented through headphones binaural recordings were made of the stimuli from each of the loudspeakers with each of the artificial and human heads at the listener position. Headphone transfer functions were measured, from which equalisation filters were designed.

Twenty listeners (10 male and 10 female) participated. They were between 20 and 30 years old and they all had



controlled normal hearing. None of the listeners had participated in localization experiments before. Since there were 10 conditions and 10 stimuli (pink noise and speech) each listener participated in 20 sessions that were balanced by means of a Latin square design.

#### 4. Results and discussion

The errors made during the experiment have been divided into four categories. If a response is not on the same cone-of-confusion as the stimulus, it is denoted an out of cone error. When errors are made by confusing directions on the same cone, it is denoted a within cone error, except for that special “cone” formed by the median plane, in which case it is denoted a median plane error. A response given in the same direction as the stimulus, but at an incorrect distance, is denoted a distance error.

The errors for every condition are shown in Figure 7, split up into the four error categories. The results are given as the percentage of errors made with respect to the potential number of errors in each category.

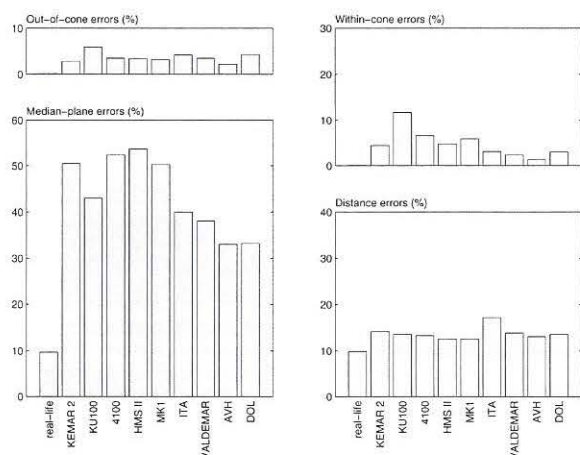


Figure 2. Error percentages divided into four categories.

In the median plane clearly the lowest number of errors occurred for real life listening followed by the two human heads. Among the artificial heads VALDEMAR and ITA gave the lowest number of median plane errors, followed by KU100. Notice, though, that KU100 shows the largest number of out of cone and within cone errors. This observation is not surprising, since it is the only head without shoulders and torso. Distance errors are remarkably similar independent of the condition, real life included. This suggests that the directional filtering of the recording head is less critical for the perception of distance, and that other cues play an important role.

#### 5. Comparison of five experiments

The experiment reported in this article is one in a series of experiments employing the same loudspeaker setup and general procedure. Although small procedural differences exist and different number of listeners participated in the individual experiments, it is justified to compare the results of 5 experiments.

##### 5.1. Experiment A

This experiment was reported in Møller et al. [2] as Experiment B. Recordings were made with the build-in microphones of 8 artificial heads and 8 listeners participated. The heads were: KEMAR 2 A, KU80 A, KU81 A, HMS I A, HMS II A, 5930 A, 4128 A and TORONTO A. The heads are named here as in the original publication and the letter A is added to denote Experiment A. Please refer to the original publication for a description of the heads and further details.

##### 5.2. Experiment B

This experiment was reported in Møller et al. [2] as Experiment C. The 20 listeners included the 8 listeners of Experiment A above and 10 artificial heads were used. Recordings were made with microphones at the blocked entrances to the ear canals of the artificial heads. The heads were: KEMAR 1 B, KEMAR 2 B, KEMAR 3 B, KEMAR 4 B, KU80 B, KU81 B, HMS I B, HMS II B, 4128 B and TORONTO B. As before the last letter, B here, denotes the experiment.

##### 5.3. Experiment C

Experiment C was done with a new group of 8 listeners and included 2 artificial heads: KU100 C and 4100 C. The KU100 artificial head was placed on a stand, and the 4100 head was used with its jacket and the neck ring in the 'forward' position. In both cases built-in microphones were employed. This experiment was carried out by Clemen Boje Jensen and not publicly reported before, but the procedures used were identical to those described in Møller et al. [2].

##### 5.4. Experiment D

This experiment was reported by Sandvad et al. [9] as Experiment B. Recordings were made with the build-in microphones of the 4 artificial heads: KU100 D, VALDEMAR D, ITA D and MK1 D. The experiment was done with 12 listeners. The listeners in Experiment D had not participated in any of the previous experiments.

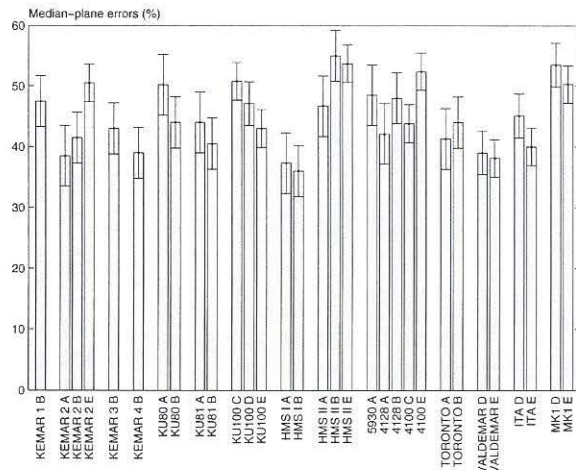
##### 5.5. Experiment E

Experiment E refers to the main experiment reported in this article. The artificial heads are named as follows: KEMAR 2 E, KU100 E, HMS II E, 4100 E,

VALDEMARE, ITA E and MK1 E. The 20 listeners in this experiment were again newly recruited.

## 5.6. Results and discussion

Median plane errors for the 5 experiments are shown in Figure 10, grouped by recording head. The bars indicate the observed means, and the (small I shaped) error bars indicate 84% confidence intervals calculated for each experiment. The means of two heads with non-overlapping confidence intervals are significantly different at the 5 % level in a t-test.



**Figure 3.** Comparison of artificial heads in 5 listening experiments.

The confidence intervals were calculated from the common variances in the analyses of variance, including the random listener factor, thus making the confidence intervals valid for a population. The "odd" size of the confidence interval has been chosen so that the means of two artificial heads with non overlapping confidence intervals are significantly different at the 5% level in a t test. (To be exact, this requires a few preconditions, e.g. that the confidence intervals are of equal size for the conditions compared; the true significance level, though, will only deviate marginally from 5% for the small violations in the present material).

Systematic differences between the 5 experiments were inevitably introduced by changing details of the psychometric procedure, equipment, experimenter, group of listeners, the group's experience etc. Despite this, the confidence intervals generally overlap for a head in different experiments. However, in quite many cases the confidence intervals of different recordings heads do not overlap, indicating that the heads are significantly different from each other.

## 6. Conclusions

The experiment shows that localization with non-individual binaural recordings is generally much poorer than in real life. An analysis of variance revealed that there were no significant differences between heads for distance errors. The differences between recording heads were statistically significant for median plane errors, though, with the two selected human heads producing less median plane errors than any of the artificial heads. The accumulated results of the five experiments show that there are significant differences between currently available artificial heads.

## 7. References

- [1] H. Møller, M. F. Sørensen, C. B. Jensen, D. Hammershøi, "Binaural technique: Do we need individual recordings?", *J. Audio Eng. Soc.*, Vol. 44, No. 6, pp. 451-469, 1996.
- [2] H. Møller, D. Hammershøi, C. B. Jensen, M. F. Sørensen, "Evaluation of artificial heads in listening tests", *J. Audio Eng. Soc.*, Vol. 47, No. 3, pp. 83-100, 1999.
- [3] H. Møller, C. B. Jensen, D. Hammershøi, M. F. Sørensen, "Using a typical human subject for binaural recording", *Presented at the 100th Convention of the Audio Engineering Society*, Copenhagen, Denmark, preprint 4157, 1996.
- [4] P. Minnaar, S. K. Olesen, F. Christensen, H. Møller, "Localization with binaural recordings from human and artificial heads", *J. Audio Eng. Soc.*, Vol. 49, No. 5, pp. 323-336, 2001.
- [5] IEC 711 Standard, "Occluded-ear simulator for the measurement of earphones coupled to the ear by ear inserts", *1st ed. International Electromechanical Commission*, Geneva, Switzerland, 1981.
- [6] ANSI S3.25-1979 (R1986), "American national standard for an occluded ear simulator", *American National Standards Institute*, New York, 1980.
- [7] IEC 959 Technical Report, "Provisional head and torso simulator for acoustic measurements on air conduction hearing aids", *1st ed. International Electromechanical Commission*, Geneva, Switzerland, 1990.
- [8] ITU-T P.58 (08/96), Recommendation, "Head and torso simulator for telephonometry", *International Telecommunication Union*, 1996.
- [9] J. Sandvad, F. Christensen, S. K. Olesen, H. Møller, "Localization with artificial head recordings", *Presented at the 134th Meeting of Acoustical Society of America*, San Diego, USA, 1997. Abstract in *J. Acoust. Soc. Am.* Vol. 102, No. 5, Pt. 2, p. 3116 (paper 2pSP4).